April 2009

A Novel Approach to S. cerevisiae Metabolic Engineering for Bioethanol Production

Nicholas Paul Dufour
Worcester Polytechnic Institute

Project Number RP1 MQP 2183

A Novel Approach to *S. cerevisiae* Metabolic Engineering for Bioethanol Production

A Major Qualifying Project Report:

Submitted to the Faculty

Of the

WORCESTER POLYTECHNIC INSTITUTE

In partial fulfillment of the requirements for the

Degree of Bachelor of Science

By

_____
Nicholas Paul Dufour

Submitted: April 30, 2009

Approved:

_____
Professor Reeta Prusty Rao, Advisor

Table of Contents

Abstract

While ethanol holds promise as a fuel source, implementation is hampered due to reliance on material used in food production. Lignocellulose, an abundant and unexploited substrate, is problematic due to the inability of microbes to ferment it. Here a novel approach for *Saccharomyces cerevisiae* metabolic engineering was employed, using mutagenesis of a regulatory TATA-binding protein combined with traditional genetic engineering to produce a phenotype capable of lignocellulose metabolism. This avoids poor optimization, seen in piecewise genetic engineering, and promises a means of creating desired ethanol producing phenotypes. We report success engineering a new yeast strain, containing a desired genotype for metabolic engineering, as well as a strain that can successfully grow in a culture having only xylose and arabinose.

Introduction

Rationale

As world energy demand climbs, interest in alternative sources of fuel continues to grow. A substantial degree of research and development centers on renewable fuel sources, which are seen as a long-term solution to "peak oil" (Lovins, Datta, Bustnes, Koomey, & Glasgow, 2005)—the point at which world petroleum production peaks, and begins an invariable decline. Biofuels, fuel produced from biological material, is necessarily renewable and is the focus of intense research and development. At the forefront of biofuels research is bioethanol. Ethanol, a simple two-carbon alcohol, is advantageous as a biofuel due to the ease with which it is synthesized by microorganisms. Numerous organisms synthesize ethanol as a natural byproduct of metabolism. Brewer's yeast, *Saccharomyces cerevisiae* has been used by humans since prehistory for this very purpose (Maksoud, El Hadidi, & Amer, 1994). Ethanol is also cleaner burning than gasoline, and an ethanol-based fuel economy is expected to produce substantial increases in air quality (Sonderegger, 2004). Further, the use of bioethanol as a major source of fuel energy, particularly for automobiles, will decrease dependence on foreign oil.

Given these facts, it may appear curious that ethanol was not adopted long ago. Unfortunately, the move to an ethanol-based fuel economy is not without its own difficulties (Bastianoni & Marchettini, 1996). The greatest of these is the substrate used to produce ethanol. Large-scale microbiological production of ethanol relies on crop-based raw materials, such as corn starch and sugar cane juice (Sonderegger, 2004). These materials are under simultaneous consumption by various sectors in agriculture, where it is used to create feed, both for livestock and for human consumption. Competition from agriculture affects not only the price of food , but is calculated to account for 40% of the cost of ethanol based biofuels (Zaldivar, 2001).

Both agriculture and microbiological ethanol production use such substrates because they contain high concentrations of glucose. Glucose, a monosaccharide aldohexose sugar, is highly energetic and forms the cornerstone for the energy metabolism of most organisms on the planet. In agriculture, this energy is used by livestock. In ethanol production, the energetic molecules are converted microbially into ethanol, which combusts readily yielding the energy to perform work. Fortunately, the sugar content of crops like corn is not exclusively glucose. Other sugars, particularly the pentose sugars xylose and arabinose, exist in abundance. In plants, xylose and arabinose, along with glucose, are localized in the polymer complex lignocellulose. Lignocellulose is largely unused by humans and as livestock feed, since digestion of the tough polymers is nearly impossible.

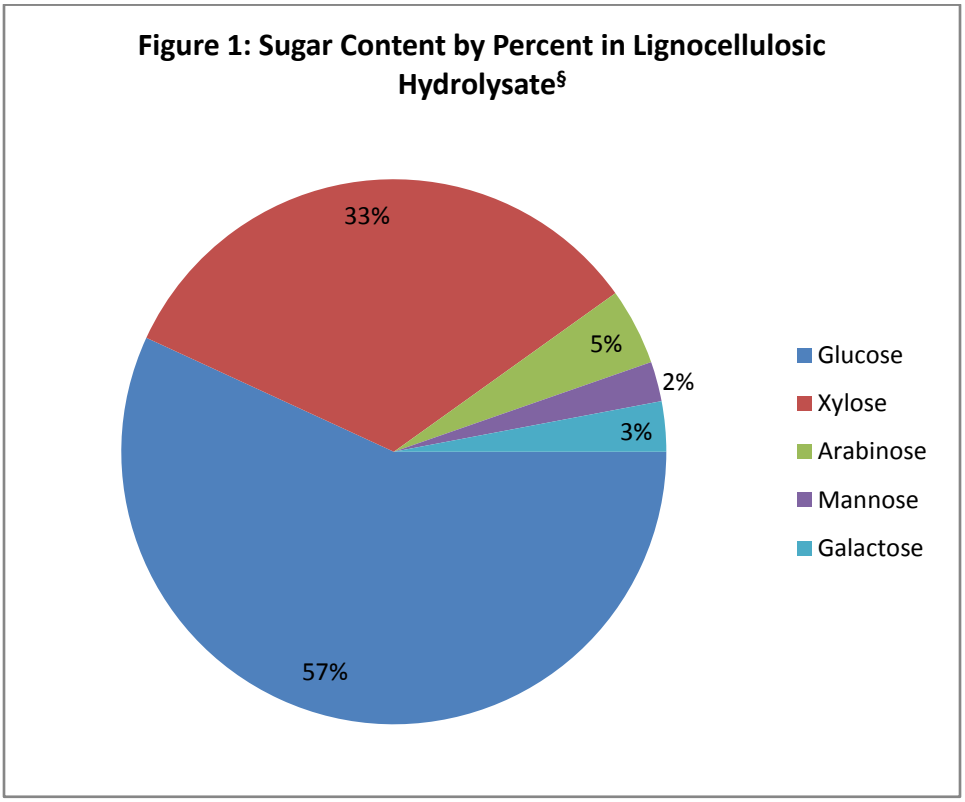


Figure 1: Sugar content by percent in lignocellulosic hydrolysate.
§Adapted from data: (Tkáč, Gemeiner, Švitel, & Benikovský, 2000).

Lignocellulose is actually composed of three polymers: lignin, hemicellulose, and cellulose. Lignocellulose constitutes the bulk of plant cell walls (Sun & Cheng, 2002), and forms the basis of biomaterials like wood. It is present in nearly all plants and is astonishingly abundant. It is calculated to be the most abundant polymer in nature (Watanabe, 2007), and accounts for more than half the biomass on the planet (Zaldivar, 2001). Since it is not easily used by the agricultural industry, it is discarded and forms the primary constituent of many agricultural waste streams (Sun & Cheng, 2002). Table 1 details the lignocellulosic content, by percentage, of some common waste products, both agricultural and otherwise.

Agricultural Waste	**Lignocellulosic Content (%)**
Wheat straw	95%
Discarded Newspapers	~100%
Leaves and Lawn Refuse	90%
Swine waste	>30%
Cattle manure	15%

Table 1: Lignocellulosic content of common biological wastes. Adapted from (Sun & Cheng, 2002).

Thus, use of lignocelluloses the primary substrate for fermentation is now seen as necessary to the establishment of viable ethanol biofuel production, specifically it is necessary for ethanol to compete viably with fossil fuels (Hinman, 1989). Currently, it is nearly completely unutilized despite being an enormous potential source of energy.

Collectively termed “lignocellulosic alcohol” (Lynd, 1991), the use of lignocelluloses to produce bioethanol (along with other alcohols, such as butanol) is problematic. No known eukaryote will grow on xylose or arabinose anaerobically, and even in bacteria the primary means of growth on either sugar is non-fermentative, instead using an aerobic metabolism (Jin, 2004) although exceptions do exist. Much effort has been exerted in engineering microbes for efficient fermentation of lignocellulosic sugars. Attempts at achieving fermentation have been

performed on a number of organisms, with most focusing on the Brewer's yeast *Saccharomyces cerevisiae*. *S. cerevisiae* is used in nearly all industrial fermentation to produce ethanol (Watanabe, 2007) (Jin & al., 2002) due to its ability to efficiently ferment glucose into ethanol and its naturally high tolerance for ethanol in solution. In spite of the discovery of an endogenous pathway for xylose metabolism in *S. cerevisiae* (Toivari, 2004), progress has been hampered due to the respiratory response exhibited by *S. cerevisiae* when exposed to xylose, resulting in growth but a distinct lack of fermentation, as most xylose is converted into xylitol (Jin, 2004). Thus, the results of previous efforts in producing a viable, industrial microbial strain for the production of ethanol from lignocelullosic sources may be described, at best, as mixed.

Project Description

A relative of *S. cerevisiae*, the yeast *Pichia stipidus*, possesses enzymes thought to be useful for xylose metabolism. This provoked much of the work in *S. cerevisiae* metabolic engineering to focus on the use of transgenes from *P. stipidus.* Transgenic modification, in parallel with aggressive overexpression of genes coding for non-oxidative pentose phosphate pathway enzymes, have been used to ameliorate the aforementioned respiratory response and minimize the effects of limiting metabolic steps (Karhumaa, 2005).

These two trends in metabolic engineering represent the two guiding principles of this MQP: novel gene introduction and metabolic optimization. Novel gene introduction took the form of a highly versatile cassette of transgenes from a variety of organisms, coding for proteins involved in metabolic transport processes and fermentative xylose and arabinose metabolism. Metabolic optimization involved targeted gene deletion, and an innovative means of eliciting global metabolic change to create novel, *de novo* phenotypes termed **g**lobal **t**ranscription **m**achinery **e**ngineering, or gTME.

Additionally, a number of accessory assays and techniques had to be developed to permit these two methods to be fully realized. The process of gTME, relying on the generation of numerous mutant clones (*see* Methods), necessitates either a method of iterative selection or of high-throughput phenotype detection, both of which were developed for the purposes of this MQP.

A workflow for successful completion of these goals was set forth ahead of time, and consisted of a number of interdependent, orderly steps. These steps are represented schematically in Figure 2.

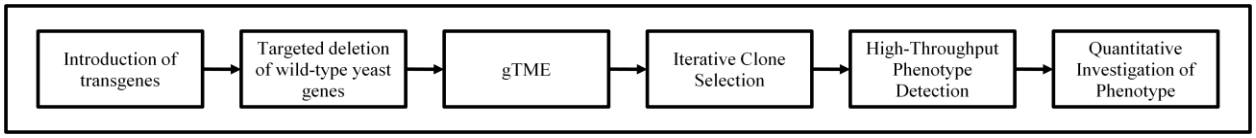


Figure 2: Summary of workflow process for the engineering of a viable, xylose and arabinose *fermenting yeast strain.*

Methods

The Lignocellulosic Transgenic Cassette

Overview

Wild-type *S. cerevisiae* is unable to ferment or utilize xylose and arabinose effectively (Hamacher, 2002). A number of other microorganisms are more readily able to utilize such pentose sugars, although not necessarily fermentatively. These microorganisms include the yeast *Pichia stipidus* (Jin, 2005), the bacterium *Escherichia coli* (Sedlak, 2001), and the bacterium *Cochliobolus carbonum* (Apel Birkhold, 1996). The basis for this difference is genetic, and a number of studies have revealed specific genes responsible for such metabolism, coding for enzymes that digest xylose and arabinose.

Based on research done on the variety of genes implicated in xylose and arabinose metabolism in other organisms, a list of genes was generated consisting of an array of nine genes coding for metabolic enzymes and sugar transport proteins and under the control of a powerful transcription elongation factor promoter (pTEF). This promoter acts to constitutively express genes under its control, hopefully resulting in high levels of gene expression. This list of genes would eventually become the Lignocellulosic Transgenic Cassette, or LTC.

The LTC was introduced to the *S. cerevisiae* genome by homologous recombination, a process where the yeast cell integrates genetic information into its genome by excising the material already present there. The location at which this occurs, and if it occurs at all, depends on the homology between the original and new sequences, as long as the flanking regions are where homology lies. The LTC is flanked by regions homologous to the gene *SPT15*, which codes for the TATA-binding protein (TBP) component of the transcription factor complex TFIID (*see* Methods, Section 3: gTME). Integration at this site disrupts the original *SPT15* in yeast,

which would result in an nonviable clone. Thus the LTC also carries another copy of *SPT15*, a mutant[1] previously generated by Alper et al, using gTME, and found to upregulate glucose fermentation (*see* Methods, Section 3: gTME) (Alper, 2006).

Additionally, the LTC contains a geneticin (G418) resistance gene ($G418^R$). G418, an antibiotic, is used as a selection mechanism and is used in eukaryotic microbiology as kanamycin or ampicillin is used in bacterial studies. The $G418^R$ is flanked by LoxP sites, permitting excision by Cre recombinase. Once removed by a Cre/LoxP system, $G418^R$ can be used again to introduce or delete other genes. In addition to G418 selection, the presence of the LTC was verified by PCR.

Genetics of the LTC

The LTC contains 9 genes that pertain directly to the xylose and arabinose fermentation. The structure of the cassette is represented schematically in Figure 3.

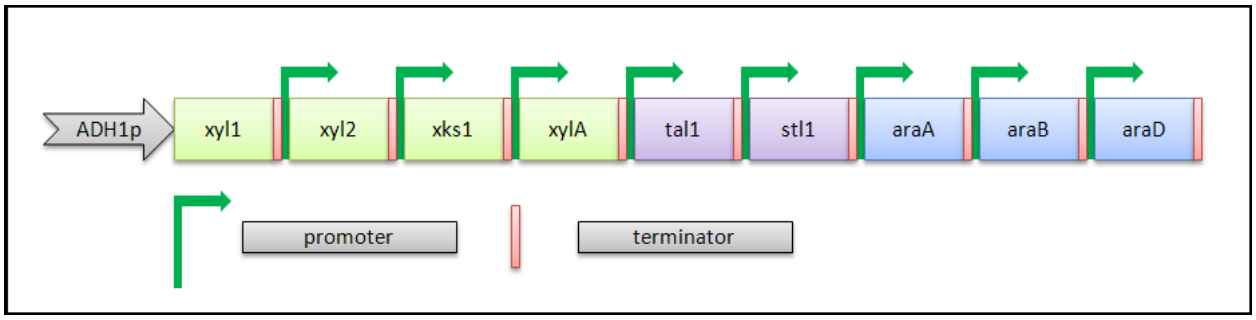


*Figure 3: A schematic representation of the structure of the lignocellulosic cassette. Green arrows denote the pTEF promoter. Note that this schematic includes only genes directly relevant to xylose/*arabinose *metabolism, and omits accessory flanking regions. The total size of the LTC is approximately seventeen thousand base pairs (17kb)*

[1] All mutant copies of *SPT15* are denoted *spt15*Δ. Thus, *spt15*Δ denotes any mutant copy or set of such copies of *SPT15* in the mutant library.

The LTC contains transgenes from 4 different organisms, as well as several genes from *S. cerevisiae* wild-type that are overexpressed through the use of the *pTEF* promoter. Table 2 details the source organism, protein, and purpose of every gene represented in Figure 3.

Gene	Source Organism	Protein	Purpose
Spt15Δ	*Saccharomyces cerevisiae*	TATA-binding protein (TBP)	This is the *SPT15* mutant shown by Alper et al to be upregulate ethanol synthesis and resistance in liquid culture.
XYL1	*Pichia stipitis*	Xylose Reductase (XR)	A xylose metabolizing enzyme found to be beneficial to ethanol production in a xylose-based culture medium (Walfridsson, Anderlund, & Bao, Expression of different levels of enzymes from the Pichia stipitis XYL1 and XYL2 genes in Saccharomyces cerevisiae and its effects on product formation during xylose utilisation, 1997), which reduces D-xylose to xylitol.
XYL2	*Pichia stipitis*	Xylulose reductase/Xylitol dehydrogenase	A xylose metabolizing enzyme that oxidizes xylitol to D-xylulose (Walfridsson, Anderlund, & Bao, Expression of different levels of enzymes from the Pichia stipitis XYL1 and XYL2 genes in Saccharomyces cerevisiae and its effects on product formation during xylose utilisation, 1997).
XYL3 (*XKS1*)	*Pichia stipitis*	Xylulokinase	A xylose metabolizing enzyme that converts D-xylulose to D-xylulose-5-P, permitting entry into the pentose phosphate pathway (Hahn-Hagerdal, 2001).
XYLA	*Streptomyces diastaticus*	Xylose Isomerase	A xylose metabolizing enzyme that converts D-xylose to D=xylulose, effectively providing an alternative to the *XYL1* → *XYL2* pathway (Gong, 1981).
TAL1	*Saccharomyces cerevisiae*	Transaldolase	An enzyme that catalyzes a non-oxidative phase of the pentose phosphate pathway. The overexpression of *TAL1* has been shown to increase flux through the pentose phosphate pathway (Walfridsson, M, X, & Hahn Hagerdal, 2001)
STL1	*Pichia stipitis*	Sugar Transporter Like protein	A protein associated with cellular transportation, shown to be necessary for xylose transport (Bolesa & Hollenberga, 2006).
ARAA	*Salmonella typhimurium*	L-arabinose isomerase	An enzyme responsible for catalyzing the conversion of L-arabinose to L-ribulose (Wisselink, et al., 2007).
ARAB	*Salmonella typhimurium*	L-ribulokinase	An enzyme responsible for catalyzing the conversion of L-ribulose to L-ribulose-5-P, having a role in the arabinose metabolic pathway, that, together with *ARAD*, is homologous to that of Xylulokinase (Wisselink, et al., 2007).
ARAD	*Bacillus subtilis*	L-ribulose-5-phosphate 4-epimerase	An enzyme responsible for catalyzing the conversion of L-ribulose-5-P to D-xylulose-5-P permitting entry into the pentose phosphate pathway (Wisselink, et al., 2007).

Table 2: Constituent genes of the LTC. The functional genes of the LTC are displayed, along with the source organism, protein coded for, and the purpose in the context of Yeast xylose/arabinose metabolism.

LTC Metabolic Pathways

The LTC is intended to introduce several new metabolic pathways in yeast. Putatively, four such pathways were introduced, two for xylose metabolism and two for arabinose metabolism. The goal of all new pathways is to generate D-xylulose-5-P, which enters the pentose phosphate pathway, generating NADH and 5C-sugars for nucleotide synthesis. The end products of the pentose phosphate pathway, fructose-6-P and glyceraldehyde-3-P, join glucose in entering glycolysis. Glycolysis produces phosphoenolpyruvic acid, converted into pyruvate by pyruvate kinase, and ultimately resulting in ethanol production.

These new metabolic capabilities are detailed in Figure 4 (introduced pathways are in dashed lines), along with the deletion (in red lines) of the *GRE3* pathway (*see* Methods Section 2: Deletion of *GRE3*).

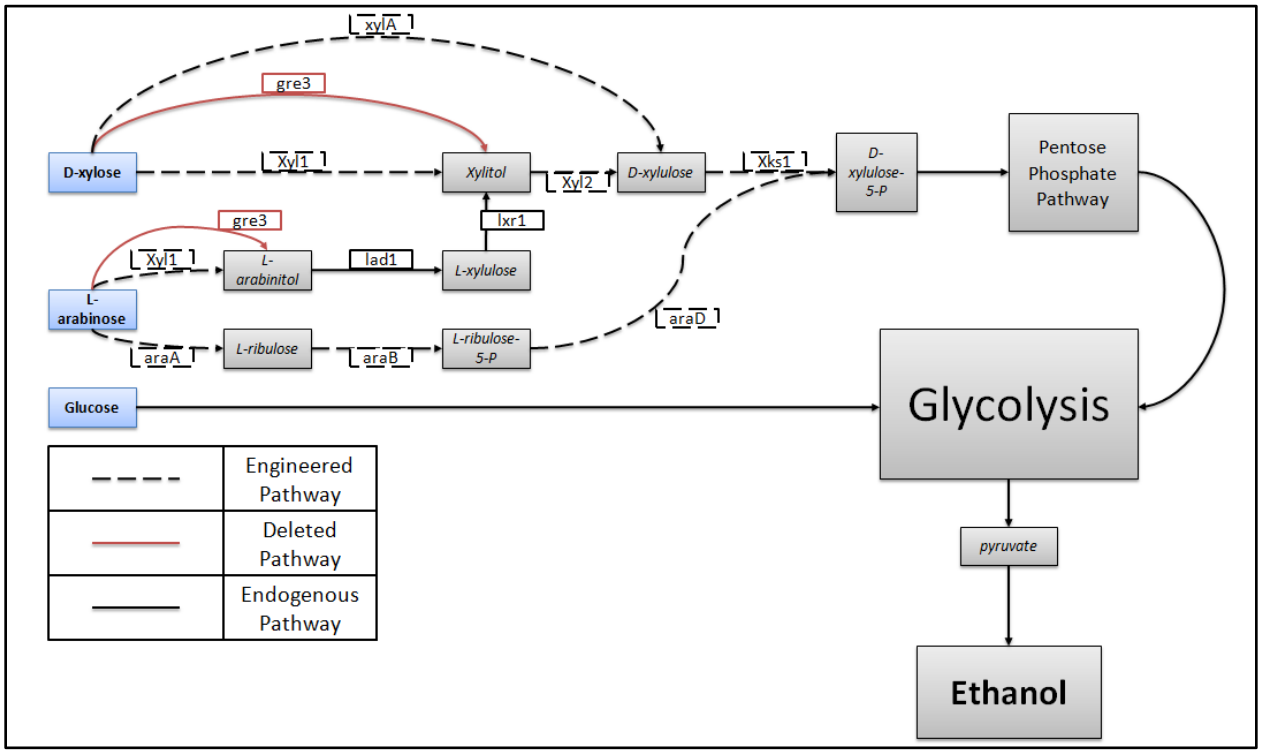


Figure 4: Metabolic pathways introduced through the LTC. Exogenous reagents (D-xylose, L-arabinose, Glucose) are shown in blue and bold. Intermediary molecules are in italics. Metabolic processes are shown in grey boxes without italics. Introduced pathways are shown in dashed lines, along with corresponding genes in dashed bounding boxes. Red lines and red bounding boxes indicates a deleted gene / pathway (*see* Methods Section 2: Deletion of *GRE3*).

Deletion of *GRE3*

GRE3 codes for an aldose reductase. This aldose reductase is nonspecific (Petrash, 2001), and accounts for the endogenous pathway for xylose and arabinose metabolism in yeast (*see* Introduction Part 1: Rationale). As seen in Figure 3, *GRE3* introduces a means to convert xylose into xylitol and arabinose into arabinitol. From there, arabinitol can be converted into xylitol by wild-type metabolism. However, it does not afford yeast the necessary steps to convert the

resulting xylitol into xylulose or D-xylulose-5-P, necessary to achieve fermentation of the lignocellulosic sugars.

The inefficacy of *GRE3* in providing efficient fermentative metabolism for xylose and arabinose is perhaps due to its utility as a stress response. *GRE3* is induced, under regulation by the **h**igh **o**smolarity **g**lycerol, or HOG, pathway. HOG is activated by deleterious changes in the extracellular environment, particularly hyperosmotic stress. The main function of HOG is to increase the concentration of glycerol in the yeast cell, mitigating the effects of the osmotic stress (Dihazi, Kessler, & Eschrich, 2004). *GRE3,* being involved in the stress response, does not participate in normal metabolism and is not part of a complete fermentative pathway.

To make matters worse, research has shown that the aldose reductase product of *GRE3* actively inhibits xylose and arabinose metabolism in transgenic yeast (Traff, 2001). To ameliorate this problem, we constructed an insert to be introduced into yeast by homologous recombination, replacing *GRE3* with $G418^R$. This was accomplished by amplifying $G418^R$ from a plasmid using primers having flanking regions homologous to the chromosomal DNA sequences immediately surrounding *GRE3*. This construct is then introduced into the chromosome, recombining at the *GRE3* site. The presence of $G418^R$ was verified using PCR, where the primers were selected that product would only be formed if (a) $G418^R$ was present in the genome and (b) the $G418^R$ resistance gene was present in the correct location.

The goal of the manipulations was to create modified yeast that, with the exception of the LTC and the deletion of *GRE3*, was as close to wild-type as possible. As such the $G418^R$ was flanked with LoxP sites, similar to the $G418^R$ in the LTC (*see* Methods Section 1: The Lignocellulosic Transgenic Cassette Part 1: Overview). This permitted the G418 resistance

cassette to be excised using Cre recombinase. The *GRE3* deletion method is represented graphically in Figure 5.

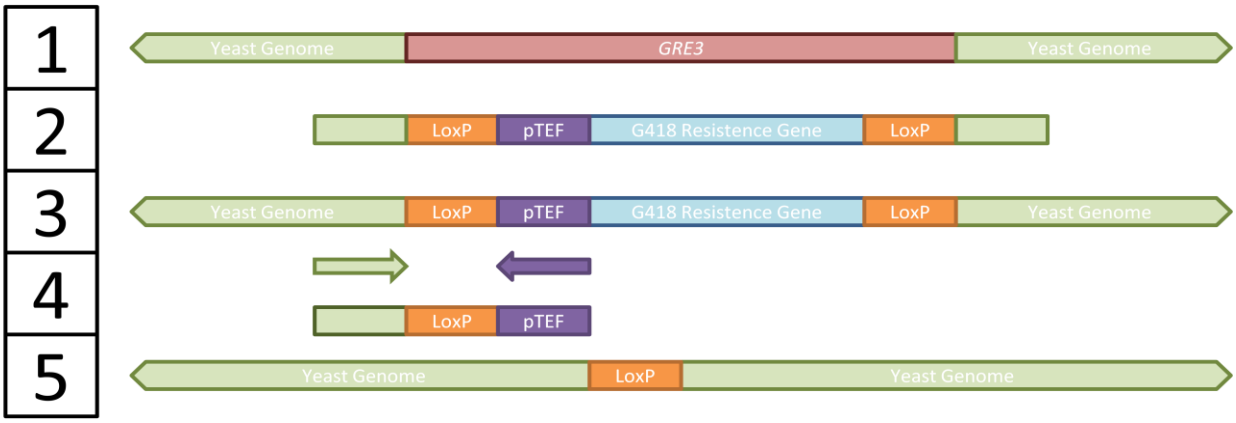


Figure 5: GRE3 Deletion Method. Figure 4 depicts the series of steps required to delete GRE3 using genetic techniques. (1) GRE3 surrounded by the yeast genome. (2) The constructed G418 deletion insert, consisting of flanking regions homologous to the yeast DNA immediately surrounding the GRE3 gene in the chromosome amplified through the use of a yeast genome homologous primer, LoxP sites, and in the center G418^R under the control of the promoter pTEF. (3) The G418 deletion insert is introduced to the yeast genome in place of GRE3 by homologous recombination with the homologous flanking regions of the deletion insert with GRE3. (4) PCR is conducted with primers (purple) to ensure that G418 is present and in the correct location. (5) The deletion insert is removed by introduction of Cre recombinase, which excises the resistance gene at the LoxP sites.

Global Transcription Machinery Engineering (gTME)

Overview

gTME is a novel technique developed by Hal Alper and Gregory Stephanopoulos to create new metabolic phenotypes in microorganisms in a rapid and profound manner (Alper, 2006). gTME was developed in response to frustration at the difficulty of metabolic and genetic engineering to produce new metabolic phenotypes that are functional and optimized. Metabolism is very complex, having a complexity density exponentially higher than its progenitor genome (Goryanin, Hodgman, & Selkov, 1999), and to simulate even the most basic bacterial metabolisms at a rate approaching real-time requires a supercomputer utilizing complex control

analysis, ordinary and partial differential equations, bifurcation analysis, and constrained optimization solvers (Goryanin, Hodgman, & Selkov, 1999). In addition, such simulations are necessarily specific to the organism of interest.

Attempting to augment a metabolism in a completely optimal manner using the introduction of transgenic genes and modifications to the host genome would requires prohibitively large computational capacity. Thus we currently settle for brutish methods of bludgeoning a genome into creating a phenotype that is *something like* what is desired through the introduction of one gene at a time, a process both expensive and lengthy. This is generally performed using piecewise genetic modification, in a manner akin to the aforementioned genetic modification (*see* Methods Section 1: The Lignocellulosic Transgenic Cassette, *see* Methods Section 2: Deletion of *GRE3*). We use a novel method of metabolic engineering, gTME, in conjunction with traditional metabolic engineering to circumvent the above problems to generate a phenotype that produces ethanol lignocellulosically.

gTME uses random mutagenesis of globally-acting transcription factors. These globally acting transcription factors modify gene expression genome-wide and simultaneously. When this is performed using PCR, it is possible to generate billions of copies of the transcription factor gene to be transformed into the desired organism. This results in a large number of clones, each with a potentially novel phenotype that can be evaluated.

The gTME approach has numerous advantages. gTME produced clones can be forced to undergo a selection process (*see* Methods Section 4: Iterative Clone Selection), narrowing down the large number of clones to just a small number of candidates having exceptional phenotypes—allowing the results of what is essentially millions of dice-rolls to be compressed into just a few

strains. Further, because the process acts globally, it can potentially account for the fineries of epistatic regulation, or other phenomena that occur genetically that are difficult to predict by geneticists. For instance, the metabolic pathways introduced (*see* Methods Section 1: The Lignocellulosic Transgenic Cassette Part 3: LTC Metabolic Pathways) may require a number of other genes on the yeast genome to be up-or-downregulated, perhaps in a relatively minor manner, to achieve optimal xylose and arabinose fermentation. gTME, by modification of the transcription machinery that affects gene expression, can potentially account for this.

gTME is combinatoric, relying on the law of large numbers and an enormous solution space. Therefore, each individual gTME transformant will not necessarily possess these features. Further, the investigators are not necessarily aware of the possible performance of the mutated transcription factors or—importantly—the mechanism that underlies their improved performance. In this sense, success with gTME does not require complete knowledge of the target metabolism, it is *a priori* in terms of the means through which novel phenotypes are elicited.

Process Specifics

Target Gene

The transcription machinery targeted during this investigation was the gene *SPT15*. *SPT15* codes for the TATA-binding protein (TBP) subunit of the transcription factor complexes TFIID and TFIIB (Cormack & Struhl, 1992). TBP is essential for the viability of yeast cells and plays an enormous role in gene expression. Conserved widely among all eukaryotes and some archaea, TBP directs the transcription of genes by recruiting all three polymerases (Polymerase I, II, and III) to the transcription start site (Hampsey, 1998). Additionally, much gene regulation takes

place by targeting TBP with coactivators and corepressors, which interact with TBP to augment gene transcription (Lewis & Reinberg, 2003).

Vector

SPT15 was obtained cloned in a Ura3 marked plasmid vector, pRS316. The vector, known originally as pDE28-6, was generously provided by Dr. Gregory Prelich of the Albert Einstein College of Medicine of Yeshiva University. *SPT15* was excised from the vector using high-fidelity PCR amplification, with primers *SPT15*F (sense; T_m = 66.5˚C) and *SPT15*R (antisense; T_m = 62.9˚C). These primers included a SalI site and NheI site, such that the insert could be cloned back into another vector once the low-fidelity PCR was complete.

Low-fidelity PCR

Low-fidelity PCR was performed using the GeneMorph PCR mutagenesis kit from Stratagene, containing Mutazyme DNA Polymerase, an engineered DNA polymerase prone to errors. This polymerase was selected based on its ability to produce all possible transition and transversion mutations with equal likelihood.

To perform the low-fidelity PCR, *SPT15* produced from pRS316 using high-fidelity PCR was diluted to a concentration of 5 ng/µL (in general, approximately a 1:3000 dilution from the original solution). The reaction mixture is detailed in Table 3.

Reagent	Amount
rdH_2O	40μL
SPT15 PCR product (1:3000)	1μL
*SPT15*F primer	1μL
*SPT15*R primer	1μL
10x Mutazyme Reaction Buffer	5μL
40mM dNTPs	1μL
Mutazyme Polymerase	1μL
Total Volume	50μL

Table 3: Low-fidelity PCR reaction mixture. Table 3 details the reaction used for the low-fidelity Mutazyme DNA polymerase PCR reactions that generated mutant SPT15. Note that, as long as the ratios are preserved, success has been achieved scaling the above reaction up.

The PCR reaction was performed on a thermal cycler. The specific thermal program used was in accordance with the guidelines for using Mutazyme DNA polyermase. The thermal program used is detailed in Table 4.

Temperature	Time (min:sec)	Repeats
95.0˚C	02:00	1x
95.0˚C	00:30	30x
61.5˚C	00:30	
72.0˚C	01:00	
72.0˚C	10:00	1x
4.0˚C	∞	1x

Table 4: Low-fidelity PCR thermal program. Table 4 details the thermal program used for the low-fidelity PCR. The total runtime for the program is approximately 1 hour, 30 minutes. Note that ∞ denotes an indefinite period.

In most cases, low-fidelity PCR was performed in rounds. The first low-fidelity PCR reaction was stored, and a small sample was used as the template for the next round of low-fidelity PCR reaction. Repeating the reaction takes already mutated versions of *SPT15* and mutates them further—introducing new mutations in the DNA in addition to previous mutations. Typically 3 rounds of low-fidelity PCR were conducted, resolved on an agarose gel, then the 1^{st}, 2^{nd}, and 3^{rd} round reaction solutions were either ligated into a plasmid vector (pRS416) and cloned into *E. coli* or used to transform yeast in parallel.

Cloning

Cloning was conducted in two ways. Each cloning method contains different advantages and disadvantages. Succinctly, *Escherichia coli* mediated plasmid amplification permits the easy creation of clonal populations and an ease of extracting *spt15*Δ in the event the resulting phenotype is found to be effective. In addition, since there is already a copy of *SPT15* on the genome[2], it allows for automatic verification that the new *spt15*Δ is dominant.

Direct yeast transformation by homologous recombination removes the added step of *E. coli* mediated plasmid amplification, and thus can potentially generate a much larger set of clones since the digestion/ligation process of constructing an *spt15*Δ vector library necessarily cuts out the majority of *spt15*Δ molecules from the low-fidelity PCR. There are drawbacks, however, in that the yeast population is not clonal but rather consists of a number of cells having unique copies of *spt15*Δ. Further, because the *spt15*Δ molecule overwrites the previous *SPT15* copy, it is impossible to check if it is dominant directly. The *spt15*Δ introduced is also more difficult to extract than if it were on a plasmid.

Cloning the *spt15*Δ pool in *E. coli*

In this method, *spt15*Δ library is digested with the restriction enzymes SalI and NheI. The restriction sites were introduced by the primers used, and should avoid mutation by the low-fidelity DNA polymerase. This reaction is performed sequentially, first using NheI, then (in the same solution) the pH is augmented and SalI is added. The vector, pRS416, is a plasmid containing the selective markers Ura3 and Amp^R. pR4316 is digested with XbaI and XhoI simultaneously. The *spt15*Δ and pRS416 digestions are then ligated together, forming the vector pRSspt15μ.

[2] The LTC contains a mutant copy of *SPT15*, generated by performing gTME, and found by Alper et al to increase ethanol tolerance and production. It is dominant over the wild-type *SPT15*, and as such *spt15*Δ molecules introduced will be dominant over wild-type *SPT15* if they are dominant over the mutant copy *SPT15*, assuming this form of genetic transitivity holds.

Using Invitrogen TOP10 Subcloning Efficiency *E. coli*, pRSspt15µ is cloned into *E. coli* and selected for on Ampicillin-containing LB plates. Colonies are isolated, and expanded. A plasmid preparation is then conducted using Qiagen Miniprep kits. The purified, clonal pRSspt15µ copies are transformed into yeast using standard procedure.

Direct Yeast Transformation by Homologous Recombination

By omitting the *E. coli* mediation, this method is more rapid and generates more potential clones. In lieu of restriction/ligation, the low-fidelity PCR products are purified using a Qiagen PCR Purification kit, then cloned directly into yeast using standard procedure. Using this method, an aggressive selection mechanism is necessary (*see* Methods Part 4: Iterative Clone Selection).

Iterative Clone Selection

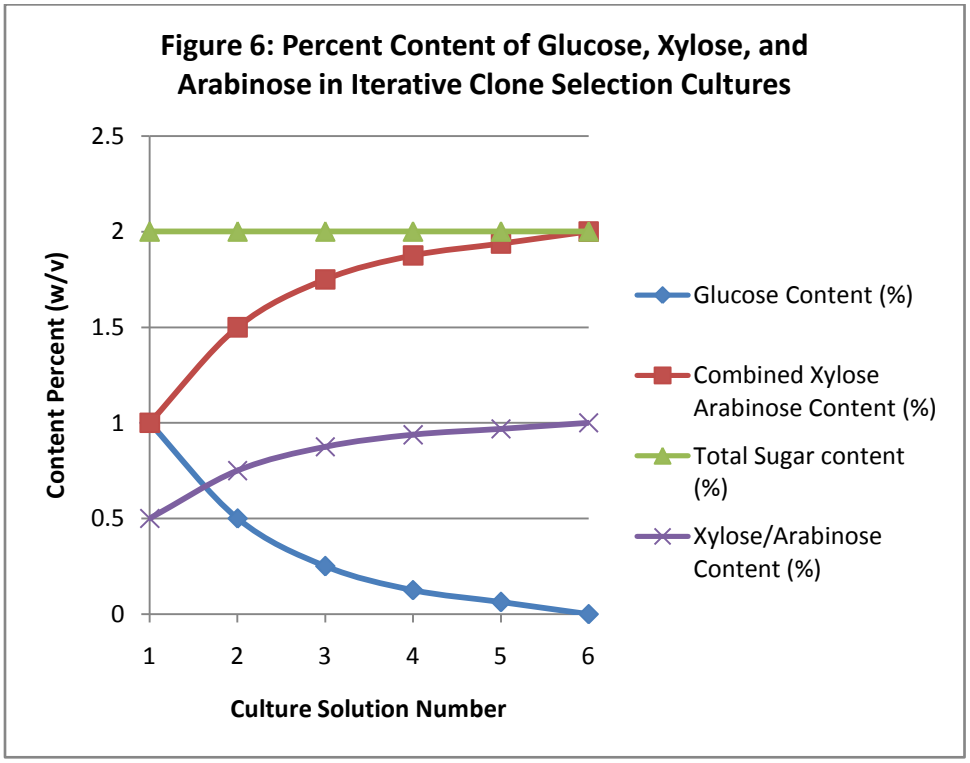
Because gTME creates such a large number of clones, going through by hand and testing them individually is humanly infeasible. A method of iterative selection was adopted, by using repeated subculturing in a variety of cultures. Six solutions of standard yeast growth media, yeast peptone or YP (Sherman, Fink, & Hicks, 1986) were produced, each having a differential amount of glucose, xylose and arabinose. The total amount of all three sugars, combined, was always equal to 2% (w/v). The content of xylose and arabinose were equal, and together summed to the difference between 2% and the total percentage of glucose present in the solution.

Beginning at 1% glucose, the amount of glucose decreased by half the previous value in each new culture, causing the combined xylose and arabinose percent content to increase by exactly the same amount to maintain the sum of 2% total. Table 5 details the concentration of each sugar by the culture solution of YP.

Solution Number	Glucose Concentration (w/v)	Combined Xylose and Arabinose Concentration (w/v)	Xylose Concentration (w/v)	Arabinose concentration (w/v)
1	1.000%	1.000%	0.500%	0.500%
2	0.500%	1.500%	0.750%	0.750%
3	0.250%	1.750%	0.875%	0.875%
4	0.125%	1.875%	0.938%	0.938%
5	0.063%	1.937%	0.969%	0.969%
6	0.000%	2.000%	1.000%	1.000%

Table 5: Total Sugar Content by Molecule for Each Culture Solution.

Table 5 is represented graphically in Figure 6. Figure 6 exemplifies the logarithmic nature of the falloff of glucose and the increase in xylose/arabinose content, which permits initial recovery following the transformation followed by rapidly more intensive selection.



Yeast cells were cultured, following transformation, in the first selective solution (culture solution number 1) overnight at 30˚C, and then aliquoted into the next culture solution, cultured overnight, and so on. Ultimately, the last culture solution contains no glucose—hence any yeast found growing in it should be able to utilize xylose and arabinose effectively, depending on the change in optical density. The last serial culturing was repeated over several days to ensure that no residual glucose remains from any previous glucose-containing cultures.

Phenotype Detection

Two distinct forms of phenotype detection were employed. Both relied on measuring the amount of ethanol being produced by the yeast culture. The first method, involving spectrophotometry, was reliably quantitative and could be related to exact amounts of ethanol in solution. The second method, using agar plates, was high throughput but would introduce errors with a higher likelihood. Despite the similarity in the assays, in terms of their quantitative value, they assay subtly different characteristics, and will be used in parallel. The spectrophotometric assay, measuring the quantity of ethanol in the culture at a given time point, quantifies that amount of ethanol produced, and, tacitly, the maximum production capacity of the yeast strain undergoing testing. Conversely, the plate assay, which utilizes strain colonies growing on agar, does not retain ethanol to the same degree and will therefore be primarily a means to measure the rate at which the yeast is producing ethanol in an initially ethanol-free environment.

The Spectrophotometric Assay

Using a kit produced by R-Biopharm, we measured the ethanol concentration from numerous cultures. The assay is based on the principle on the oxidation of ethanol by nicotinamide-adenine dinucleotide, NAD, with alcohol dehydrogenase, to acetaldehyde. Then, using low pH conditions, acetic acid is produced in the presence of aldehyde dehydrogenase. The reaction can be represented using a series of biochemical equations:

$$(1)\ Ethanol + NAD^+ \xleftrightarrow{Alcohol\ dehydrogenase} Acetaldehyde + NADH + H^+$$

$$(2)\ Acetaldehyde + NAD^+ + H_2O \xleftrightarrow{Aldehyde\ dehydrogenase} Acetic\ acid + NADH + H^+$$

The absorption is measured at 365nm after reaction (1) has been completed (the initial absorbance, A_1), and then again after reaction (2) finishes (the final absorbance, A_2). The concentration, in grams per liter can be computed:

$$c = \frac{V_T * MW}{\varepsilon * d * V_S * 2000} * F * \Delta A$$

Where

c = concentration of ethanol $\frac{grams}{liter}$
V_t = total reaction volume = 1.575 mL
MW = molecular weight of target = 46.07 $\frac{g}{mol}$
ε = extinction coefficient of NADH = 3.4 $\frac{liters}{mol * cm}$
V_s = volume of sample added = 0.050 mL
F = dilution factor (variable)
ΔA = $(A_2 - A_1)_{sample} - (A_2 - A_1)_{blank}$
d = light path, cm

The Plate Assay

In addition to the spectrophotometric assay, a high throughput method using agar plates (Jacobs & Prior, 1983) was modified from a procedure developed in 1983. Colonies were grown in liquid media having any carbon source—xylose, arabinose, or other sugars—overnight to an optical density of 0.05 at 600nm. Yeast cultures were plated onto the agar plates having a similar carbon source as the liquid culture (approx. 1uL of culture) with multiple cultures capable of being supported on a single agar plate. Another layer of agar, made in bulk, is added containing 0.5 mM 2,6-dicholorphenolindophenol (DCPIP), 3.0 mM NAD, and 9000U of Yeast Alcohol Dehydrogenase. The resulting plate is incubated at room temperature for 30 minutes. The DCPIP causes the plate to appear a deep blue. At the end of the incubation period, the indicator is added, 3mL of 0.005 M 5-methylphenazinium methylsulfate (MPMS). This results in a colorimetric

change taking place, forming a circle of yellow around the yeast colonies in direct proportion to the amount of ethanol being produced. This procedure was tested using Whatman filter paper disks saturated with solutions of distilled water and ethanol. Figure 7 is a schematic of this process.

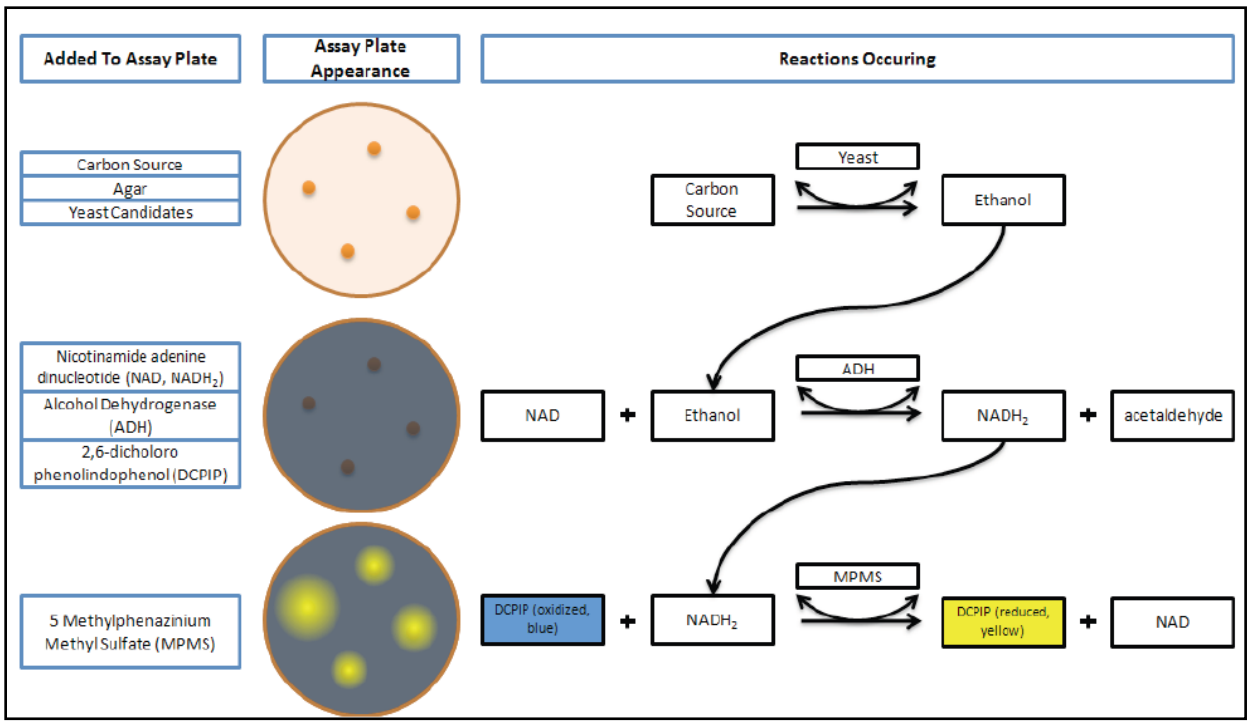


Figure 7: Graphical representation of the plate assay. Note that in the dots present on the plate represent yeast colonies. In the reactions occurring section, curved arrows connecting different reaction steps indicate that the reagent is present from the previous reaction step.

Quantitative Reverse Transcriptase PCR (qRT-PCR)

Preliminary quantitative reverse transcriptase PCR was performed on strains containing the LTC, to measure expression of LTC genes relative to the actin gene, *ACT1*, which is constitutively expressed in yeast (and in all eukaryotes). Genetic material is transmitted by messenger RNA, or mRNA. Because the mRNA is carried to the site of protein synthesis, the amount of a given protein being translated or the number of times a given gene is being transcribed is in correspondence with the appropriate mRNA molecule. A reverse transcriptase

enzyme was employed to produce complimentary DNA (cDNA) from the mRNA sample extracted from yeast using a routine yeast RNA extraction. The cDNA corresponds to the gene that was the source of the mRNA.

For the preliminary qRT-PCR, primers were selected that correspond to an enzyme from each of the three major "groups" of genes on the LTC: for the transport proteins *STL1* was chosen, for xylose metabolism *XKS1* was selected, and for arabinose metabolism, *ARAD* was selected. These were run with *ACT1*, in duplicate, as well as water as a negative control in a quantitative PCR thermal cycler. The topmost row was run using standard dilution, the next row down, all samples were diluted 1-to-10, and so on to the sixth row, diluted by a factor of 10^6. Detection was achieved using the fluorophore SYBR green.

Results

Ethanol Production in Common Yeast Strains

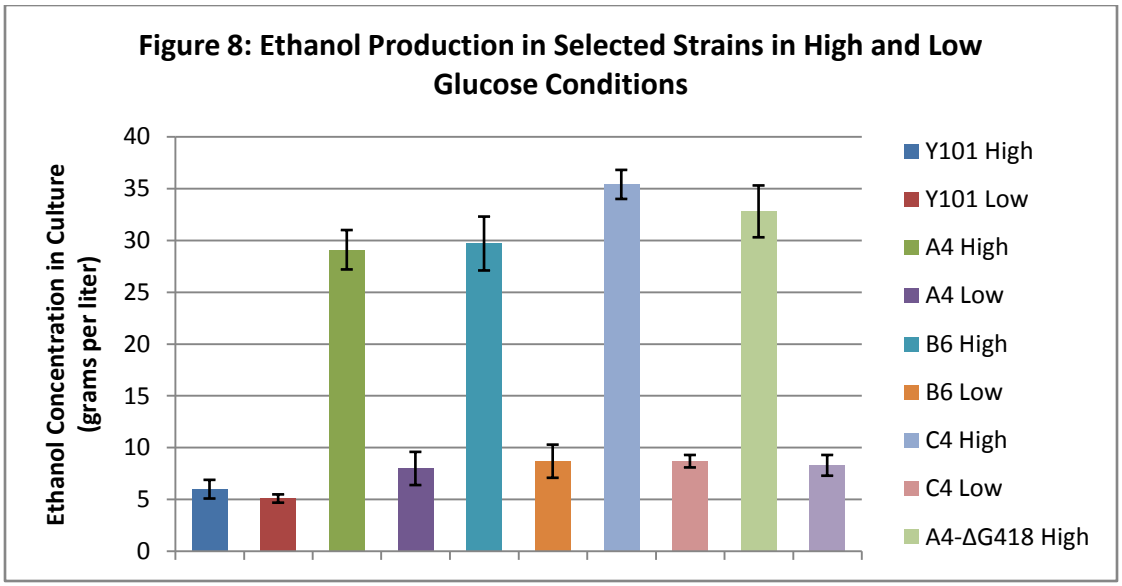
To obtain baseline data, the spectrophotometric ethanol assay was conducted on laboratory and feral yeast strains as well as on strains genetically modified in previous experiments. Strains were grown in both high glucose (10%) and low glucose (2%) cultures overnight, after which the ethanol concentrations of the resulting cultures was measured.

Strains used include the wild-type strain Y101, A4 and B6, laboratory yeast strains modified with the LTC and $G418^R$. In Table 6, data is separated by strain, and further subdivided by glucose condition. Reported is the absorbance of the culture at 600nm before being separated centrifugally, as well as initial absorption at 365nm of the culture supernatant (A_1), the post-reaction absorption at 365nm (A_2), the difference in these values (ΔA), and the calculated ethanol concentration is represented in grams per liter. Ethanol concentration was calculated in accordance with the above equation (*see* Methods Section 5: Phenotype Detection Part 2: The Spectrophotometric Assay).

Strain	Glucose Condition	Optical Density (600nm)	A1	A2	ΔA	EtOH Concentration (g/L)	Average	Standard Deviation
Y101	High	0.940	0.169	0.320	0.097	6.2	6.0	0.9
	High	0.926	0.168	0.369	0.105	6.7		
	High	0.850	0.181	0.359	0.078	5.0		
	Low	1.019	0.193	0.325	0.078	5.0	5.1	0.4
	Low	1.013	0.167	0.353	0.086	5.5		
	Low	1.009	0.166	0.340	0.074	4.7		
A4	High	2.231	0.176	0.704	0.474	30.4	29.1	1.9
	High	2.223	0.169	0.689	0.420	26.9		
	High	2.219	0.187	0.757	0.470	30.1		
	Low	1.645	0.176	0.384	0.154	9.9	8.0	1.6
	Low	1.650	0.180	0.396	0.116	7.4		
	Low	1.639	0.170	0.375	0.105	6.7		
B6	High	2.200	0.170	0.690	0.466	29.8	29.7	2.6
	High	2.004	0.162	0.764	0.502	32.1		
	High	2.043	0.168	0.690	0.422	27.0		
	Low	1.617	0.168	0.386	0.164	10.5	8.7	1.6
	Low	1.453	0.168	0.392	0.124	7.9		
	Low	1.487	0.172	0.391	0.119	7.6		
C4	High	2.244	0.174	0.773	0.545	34.9	35.4	1.4
	High	2.058	0.168	0.846	0.678	37.0		
	High	2.289	0.164	0.792	0.628	34.3		
	Low	1.875	0.170	0.371	0.147	9.4	8.7	0.6
	Low	1.592	0.163	0.390	0.227	8.1		
	Low	1.526	0.169	0.403	0.234	8.6		
A4-ΔG418	High	2.240	0.182	0.792	0.556	35.6	32.8	2.5
	High	2.045	0.188	0.778	0.590	31.4		
	High	2.184	0.171	0.760	0.589	31.3		
	Low	1.722	0.167	0.369	0.148	9.5	8.3	1.0
	Low	1.600	0.167	0.389	0.222	7.8		
	Low	1.495	0.166	0.382	0.216	7.6		

Table 6: Ethanol production in selected strains. The ethanol concentration of various strains of yeast in liquid cultures of high (10%) and low (2%) glucose concentrations is investigated after equal inoculations and 24 hours of growth.

Data in Table 6 are represented graphically in Figure 7. Note that error bars correspond to the standard deviation of the measurements taken.



Construction of Plasmid Library Containing *SPT15* Mutant Pool

Three rounds of low-fidelity PCR were conducted, such that an aliquot of the first round was used as the template for round 2, etc. Vector constructs of pRSspt15µ was conducted in accordance with the procedure detailed above.

The three sets of pRSspt15µ were transformed into TOP10 Subcloning Efficiency *E. coli*, and in turn each of the three sets were plated onto two ampicillin plates. Numerous colonies were observed, as reported in Table 7. Note that each transformation is named according to the vector pRSspt15µ with the suffix *n* where *n* is the round of low-fidelity PCR of origin for the spt15 molecule inserted, thus pRSspt15µ -1 is the set of vector constructs using the first round *spt15*Δ generated.

Plate Number	Construct Used	Number of Colonies
1	Water (Negative Control)	0
2	pRSspt15µ -1	256
3	pRSspt15µ -1	60
4	pRSspt15µ -2	472
5	pRSspt15µ -2	224
6	pRSspt15µ -3	272
7	pRSspt15µ -3	208

Table 7: Cloning Efficiency of E. coli transformation with complete pRSspt15µ *vectors.*

Of the colonies observed, between three and four were selected for expansion (four colonies were selected from plate 3 and plate 7, three from the remaining plates). Following expansion, a plasmid extraction was conducted to verify the presence of *spt15*Δ in the constructs via PCR. *spt15*Δ was verified to be present in all of the constructs, suggesting most of the observed colonies possessed *spt15*Δ. The gel produced from the PCR is displayed in Figure 9.

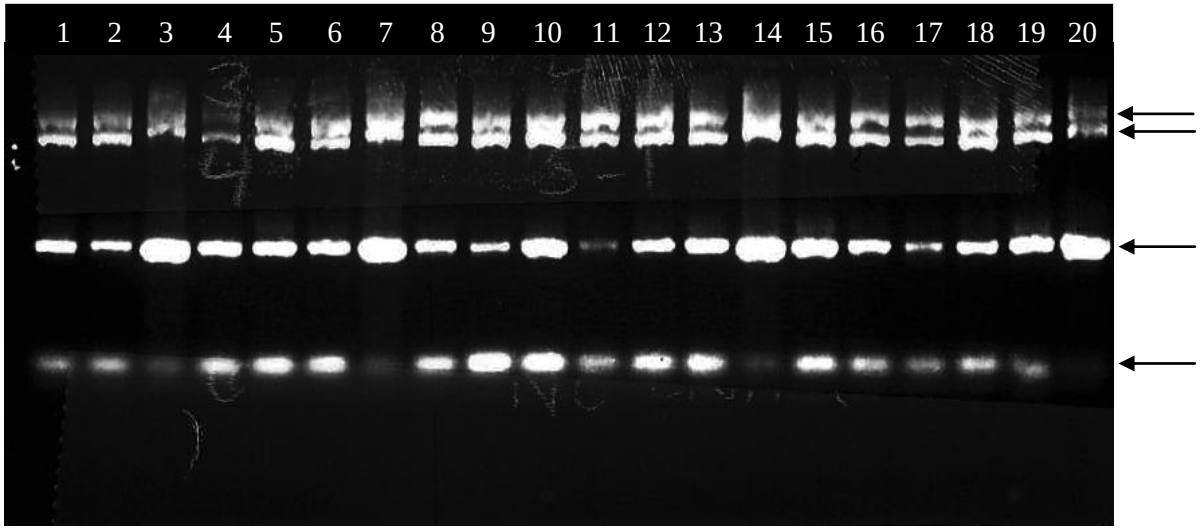


Figure 9:Gel image of SPT15 PCR on select E. coli colonies following transformation with pRSspt15µ *construct vectors. 1% agarose in TBE gel,electrophoresed at 120V for 1 hour. The topmost arrow indicates circular* pRSspt15µ, *the next is supercoiled* pRSspt15µ, *the third arrow down is spt15*Δ, *while the last arrow is primer-dimers. Lane ordering proceeds: lanes 1-3, colonies from plate 2; lanes 4-7, colonies from plate 3; lanes 8-10, colonies from plate 4; lanes 11-13, colonies from plate 5; lanes 14-16, colonies from plate 6; lanes 17-20, colonies from plate 7.*

Deletion of *GRE3*

Deletion of *GRE3* was achieved in several distinctive strains. Deletion was verified by PCR using primers as described in the Methods. Figure 10 is a scanned image of the gel produced from the PCR performed on the extracted genomic material from randomly selected colonies growing on a G418 plate. PCR was performed using the primers 5'*GRE3* and 3'*pTEF*. 5'*GRE3* primers are homologous to a portion of *GRE3* that remains intact after the deletion, while 3'*pTEF* is homologous to the *pTEF* promoter portion of the *GRE3* deletion cassette.

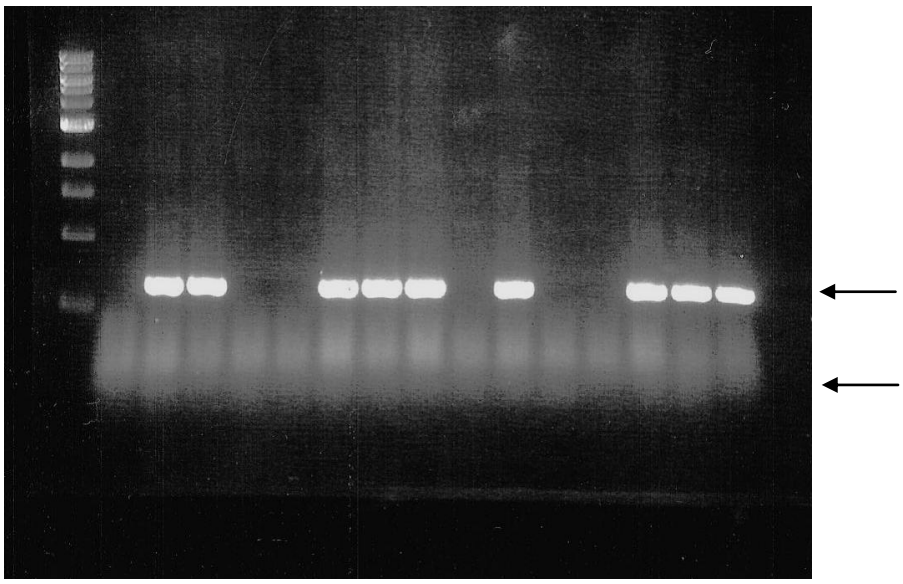


Figure10: Gel image of resolved PCR product on select the purified genomic material of strains of yeast transformed with LTC and the GRE3 deletion construct. 1% gel, electrophoresed at 120V for 1 hour. The topmost arrow indicates the product, while the second arrow indicates primer dimers. The first lane is DNA marker (NEB 1kb DNA Ladder), all the remaining lanes are randomly sampled colonies from a G418 plate. PCR product is seen in lanes 3, 4, 7, 8, 9, 11, 14, 15 and 16.

Insertion of LTC

The insertion of the LTC was the most technically difficult of the genetic manipulations, owing to its massive size (~17,000bp). The transformation efficiency was far lower than pRSspt15µ, ultimately the successful transformation produced only two candidates, which were selected on G418 YPD (yeast-peptone-dextrose) plates, both verified by PCR. The PCR was

conducted using two different sets of primers for each candidate: 5'*SPT15* and 3'*XYL1*, which should result in a product approximately one thousand one hundred base pairs (1.1kb), while the other used 5'*G418*^{R} and 3'*SPT15*, producing a product of some nine hundred base pairs (0.9kb). The use of two different sets of primers ensured the accuracy of the PCR. Figure 11 depicts the gel that was used to visualize the reaction products of the PCR that finally verified the presence of the LTC.

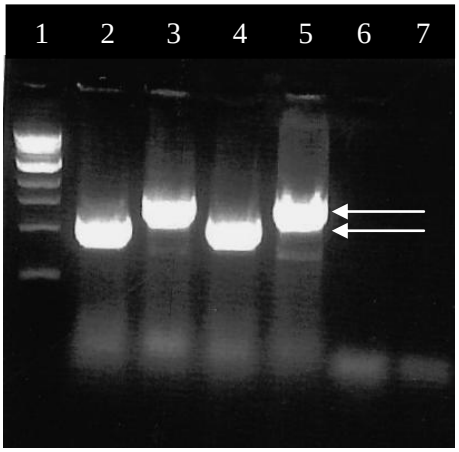
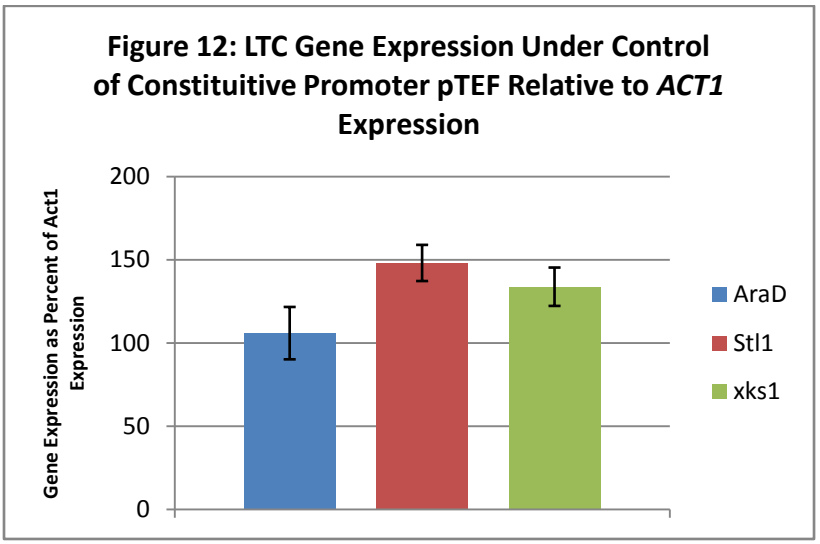


Figure 11:Gel image of PCR on verifying the presence of the LTC in the two yeast LTC candidate strains. 0.9% gel, electrophoresed at 120V for 55 minutes. The topmost arrow indicates the 5'SPT15 and 3'XYL1 PCR product, while the lower arrow indicates the 5'G418^{R} *and 3'SPT15 product. Lane ordering proceeds: Lane 1, NEB 1kB ladder; Lane 2, 5'SPT15 and 3'XYL1 PCR product for candidate 1; Lane 3, 5'G418*^{R} *and 3'SPT15 PCR product for candidate 1; Lane 4, 5'SPT15 and 3'XYL1 PCR product for candidate 2; Lane 5, 5'G418*^{R} *and 3'SPT15 PCR product for candidate 2; Lanes 6 and 7 are negative controls for both primer sets.*

LTC Gene Expression

As one of the means of novel strain characterization, quantitative reverse transcriptase PCR (qRT-PCR) was performed on *spt15*Δ and LCT transformed yeast. To achieve a baseline measurement of the LTC-transformed strains that have not yet received *spt15*Δ a preliminary qRT-PCR. Not all genes in the LTC were tested. Three were selected from each category of

genes: (1) genes that introduce xylose fermentation pathways, (2) genes that introduce arabinose fermentation pathways, and (3) genes that coded for transport proteins. As such, *XYL3*, *ARAD*, and *STL1* were selected. qRT-PCR was conducted relative to *ACT1*, a well conserved and ubiquitous gene in eukaryotes, coding for the structural protein actin. *ACT1* is advantageous for use as a relative standard because it is constitutively expressed. *ARAD*, *STL1*, and *XYL3* mRNAs were found to be 106%, 148%, and 133% as plentiful as *ACT1* mRNA. These results are depicted with standard deviation in Figure 12.



In Progress Results: A Strain Capable of Utilizing Xylose and Arabinose

Recently, a strain was produced (Y101 bearing a randomly mutated *SPT15* and the LTC) that was observed to grow in liquid cultures containing purely xylose and arabinose, with much greater efficacy than wild-type Y101. Tests on this strain, known as NDY5-3-9, or "Strain #9," are underway to investigate its capacity to grow in xylose/arabinose cultures relative to other strains, as well as its ability to produce alcohol via fermentation.

So far, experimentation has revealed a far increased capacity for aerobic growth from repeated equal cell count inoculations of Strain #9 versus wild-type Y101 in a culture of 2.6%

xylose and 2.6% arabinose, with 0% glucose. Strain #9 does not show any increased fermentative capacity over Y101 during aerobic growth ($P > 0.05$) in the same culture. Aerobic growth in xylose/arabinose for Strain #9 is almost exactly an order of magnitude greater than Y101 in equal inoculation testing ($P < 0.000001$). Strain #9 was found to grow to an optical density of 0.327, measured at 600 nm (standard deviation = 0.010), while Y101 grows to 0.37 when measured at the same optical density (standard deviation = 0.004), after 24 hours of growth with agitation at 30C. Anaerobic testing is currently in progress, as comparative growth in differential concentrations of xylose/arabinose.

In addition to anaerobic/differential lignocellulosic sugar testing, further tests will be conducted in which the yeast tested for ethanol production will be grown in cultures also containing a small amount of glucose. Research indicates that even a small amount of glucose in the culture solution (around 20mM) may induce sugar signaling cascades, improving ethanol production (Verstrepen, et al., 2004).

Discussion

The environmental relevance of the project is self-evident, due in part to global warming concerns and to the inalienable fact that fossil fuel resources are dirty and fundamentally limited. Conversion to an ethanol biofuel economy is thus a very attractive prospect: ethanol is cleaner and completely renewable. Lignocellulosic sources of ethanol production are hugely plentiful and the infrastructure for producing such sources, being essentially widely generated waste products like lawn refuse and newspapers and the byproducts of corn and paper production, is already in place. As such, if this project and projects like this achieve fruition, then such an economy would be one step closer to fruition: reliance on feedstock would be alleviated, reducing the price both of food globally and ethanol biofuels.

The difficulty in transforming the LTC was surprising, even given the known difficulty of transforming such a large cassette of genes. The primary source of this surprise was due to the fact that the success of the transformation hinged primarily on a dice-roll like random process rather than variations in the procedure. Indeed, the procedure was repeated numerous times, and it was only through replication, not necessarily variation in the procedure, that transformation was achieved.

Like the LTC, the deletion of *GRE3* was fraught with difficulty. In contrast to the experience of transforming the LTC, this difficulty was not due to the inherent difficulty of the genetic manipulation involved, but rather due to incorrect usage of the primers needed to amplify the construct used to delete it. Once the genetic manipulations were complete, testing proceeded well—both in verification (i.e., PCRs to check for the presence or absence of a specific sequence in the yeast strain) and in characterization (i.e., the qRT-PCR).

It was particularly heartening to find that the expression of LTC genes was so high, as this is predicted to simplify the possible set of mutant *SPT15* molecules that, once introduced to the yeast, would result in an optimized phenotype. This is because that, while it is possible that a mutant *SPT15* would differentially upregulate LTC genes, the probability of obtaining such a mutant that also produced additional benefits over the entire genome, for instance by adjusting epistatic regulation, is lower than obtaining a mutant that performs the latter but not the former.

The project has continued to generate excitement for all involved, due both to the relevance of the work being done to sustainability and the innovative methods employed. gTME, in particular, represents a wholly novel means of inducing phenotypic change, as well as a reversal of traditional investigative methods. The technique does not rely on prior knowledge of possible optimization mechanism, rather, the optimization is achieved first through observation of mutants and then investigation is allowed to proceed. Without methods to gather biological data in a high-throughput manner, such a technique would be impossible. However, as the means to obtain high-throughput data grows ever easier, methods like gTME will become more prevalent, allowing biology to adopt broad-based global techniques. This phase change in biology will hopefully permit an expanded understanding of the extraordinarily complex phenomena exhibited by organisms, such as metabolism.

Future Directions

The project is quite ambitious, and will be ongoing long after experiments performed explicitly for the MQP are completed. The results thus far have been largely positive, despite progress delays due to experimental difficulty and the relative complexity of the procedures. A number of goals set forth initially, and those established as the project progressed, have been met

culminating primarily in the generation of a strain of yeast having the genotype originally intended for gTME, permitting the process of gTME to begin.

There is no shortage of work remaining on the project. Some of the remaining work is hopeful, but more is absolutely necessary for the project to move forward. Since all genetic manipulations are complete and *SPT15* mutagenesis has been conducted and mutant libraries are constructed, the primary work remains in testing and quantitative procedures to gather data on the resulting mutants.

Quantitative Investigation of Phenotype

Quantitative Reverse Transcriptase PCR

One of the primary means of establishing that the LTC is responsible for enabling any novel phenotypes observed is qRT-PCR. While qRT-PCR will not be absolutely conclusive, it will show whether or not phenotypic changes correlate with changes in expression of the LTC genes or others. Although the high expression of some LTC genes has already been observed, presumably their expression will require "tweaking" to achieve an optimal phenotype. This notion is based on the presumption that, since wild-type yeast does not ferment xylose or arabinose, the genes introduced by the LTC will be of primary importance to induce such fermentation. Therefore, if changes in xylose/arabinose fermentation is observed (in the sense that gTME mutated LTC strains metabolize xylose/arabinose differentially) it will be the LTC that is responsible, at least in part.

Thus we propose to evaluate any outliers in ethanol production from the norm of LTC strains with qRT-PCR. Ultimately, the hope is that this will elucidate the proper expression ratios of LTC genes to further facilitate the desired fermentation. There is no concrete proof that any

changes will be observed, it is merely a pragmatic assumption given that either outcome will provide us with valuable data.

Gas Chromatography

Aforementioned assays will be used to evaluate ethanol production in new yeast strains generated in the project. These assays, particularly the plate assay, are not extraordinarily precise but are intended to allow rapid identification of high producing strains. Thus, as high producing strains are found, gas chromatography will be used repeatedly to establish conclusively and precisely how much ethanol is being produced to a degree that is impossible with the high throughput assays.

The length of time and the involvement required for gas chromatographic evaluation of strains may seem counter intuitive given that the principle of gTME is the production and evaluation of numerous mutant yeast constructs. However, gas chromatography will be employed after-the-fact, and used purely to verify that a subset of the strains generated which appear to be high-producing from the initial assays are *actually* high-producing.

Sequencing of pRSspt15µ for Verification of Mutation

Gene sequencing is probably the most direct approach towards verifying the efficacy of gTME in generating new phenotypes. For instance, if a new phenotype is observed in the yeast strains, but *SPT15* present in the strain (extracted by plasmid extraction or gap-repair depending on the method of *SPT15* transformation used) from the strain is wild-type or the type present already in the LTC, it is possible to conclude that this phenotypic variation is not due to gTME. If the converse is true, it will not be absolutely conclusive[3] since even given a novel *SPT15* mutant and a new phenotype, there is the danger of misattributing correlation to causation.

[3] The only absolutely conclusive method of verifying that mutant *SPT15* is having a genetic effect that has been discussed is whole-genome microarray assays, however it's likely that the cost and time involved would be prohibitive unless it is reasonably certain that a bona-fide optimized strain had been identified.

Additionally, attempting to return a sample of the yeast strain to one possessing wild-type *SPT15* could yield additional information given the possibility that it may or may not return to the previous phenotype.

Works Cited

Alper, H. (2006). Engineering Yeast Transcription Machinery for Improved Ethanol Tolerance and Production. *Science , 148*, 1565 - 1568.

Apel Birkhold, P. (1996). Cloning, Disruption, and Expression of Two Endo-β1,4-*XYLA*nase Genes, *XYL2* and *XYL3*. *Applied and Environmental Microbiology , 62*, 4129 - 4135.

Bolesa, E., & Hollenberga, C. P. (2006). The molecular genetics of hexose transport in yeasts. *FEMS Microbiology Reviews , 21*, 85 - 111.

Cormack, B., & Struhl, K. (1992). The TATA-binding protein is required for transcription by all three nuclear RNA polymerases in yeast cells. *Cell , 69*, 685 - 696.

Expression of different levels of enzymes from the Pichia stipitis *XYL1* and *XYL2* genes in Saccharomyces cerevisiae and its effects on product formation during xylose utilisation. (n.d.).

Gong, C.-S. (1981). Production of Ethanol from D-Xylose by Using D-Xylose Isomerase and Yeasts. *Applied Environmental Microbiology , 41*, 430 - 436.

Goryanin, I., Hodgman, T., & Selkov, E. (1999). Mathematical simulation and analysis of cellular metabolism and regulation. *Bioinformatics , 15*, 749 - 758.

Hahn-Hagerdal, B. (2001). Metabolic Engineering of Saccharomyces cerevisiae for Xylose Utilization. *Advancers in Biochemical Engineering/Biotechnology , 71*, 53 - 84.

Hamacher, T. (2002). Characterization of the xylose-transporting properties of yeast hexose transporters and their influence on xylose utilization. *Microbiology , 148*, 2783 - 2788.

Hampsey, M. (1998). Molecular genetics of the RNA polymerase II general transcriptional machinery. *Microbiology and Molecular Biology Reviews , 310*, 1 - 26.

Hinman, N. D. (1989). Xylose fermentation-an economic analysis. *Applied Biochemistry and Biotechnology , 20*, 391 - 402.

Jacobs, C. J., & Prior, B. A. (1983). A rapid screening method to detect ethanol production by microorganisms. *Journal of Microbiological Methods , 1*, 339 - 342.

Jin, Y.-S. (2005). Improvement of Xylose Uptake and Ethanol Production in Recombinant Saccharomyces cerevisiae through and Inverse Metabolic Engineering Approach. *Applied and Environmental Microbiology , 71*, 8249 - 8256.

Jin, Y.-S. (2004). Saccharomyces cerevisiae Engineered for Xylose Metabolism Exhibits a Respiratory Response. *Applied Environmental Microbiology , 70*, 6816 - 6825.

Jin, Y.-S., & al., e. (2002). Optimal Growth and Ethanol Production from Xylose by Recombinant Saccharomyces cerevisiae Require Moderate D-Xylulokinase Activity. *Applied and Environmental Microbiology , 69*, 495 - 503.

Karhumaa, K. (2005). Investigation of limiting metabolic steps in the utilization of xylose by recombinant Saccharomyces cerevisiae using metabolic engineering. *Yeast , 22*, 359 - 368.

Lewis, B., & Reinberg, D. (2003). The mediator coactivator complex: functional and physical roles in transcriptional regulation. *Journal of Cell Science , 116*, 3667 - 3675.

Lynd, L. R. (1991). Fuel Ethanol from Cellulosic Biomass. *Science , 251*, 1318 - 1323.

Petrash, J. (2001). Functional genomic studies of aldo-keto reductases. *Chemico-Biological Interactions , 130*, 673 - 683.

Sedlak, M. (2001). Expression of E. coli *ARAB*AD Operon Encoding Enzymes for Metabolizing L-Arabinose. *Enzyme and Microbial Technology , 28*, 16 - 24.

Sonderegger, M. e. (2004). Fermentation Performance of Engineered and Evolved Xylose-Fermenting Saccharomyces cerevisiae Strains. *Biotechnology and Bioengineering , 87*, 90 - 98.

Sun, Y., & Cheng, J. (2002). Hydrolysis of lignocellulosic material for ethanol production: A review. *Bioresearch Technology , 83*, 1 - 11.

Toivari, M. H. (2004). Endogenous Xylose Pathway in Saccharomyces cerevisiae. *Applied and Environmental Microbiology , 70*, 3681 - 3686.

Traff, K. (2001). Deletion of the *GRE3* aldose reductase gene and its influence on xylose metabolism in recombinant strains of Saccharomyces cerevisiae expressing the *XYLA* and *XKS1* genes. *Applied and Environmental Microbiology , 67*, 5668 - 5674.

Walfridsson, M., Anderlund, M., & Bao, X. (1997). Expression of different levels of enzymes from the Pichia stipitis *XYL1* and *XYL2* genes in Saccharomyces cerevisiae and its effects on product formation during xylose utilisation. *Applied Microbiology and Biotechnology , 48*, 218 - 224.

Walfridsson, M., M, A., X, B., & Hahn Hagerdal, B. (2001). Expression of different levels of enzymes from Pichia stipidus *XYL1* and *XYL2* genes in Saccharmyces cerevisiae and its effects on product formation during xylose utilization. *Applied Microbiology and Biotechnology , 21*, 53 - 84.

Watanabe, S. e. (2007). Ethanol production from xylose by recombinant Saccharomyces cerevisiae expressing protein engineered NADP+-dependent xylitol dehydrogenase. *Journal of Biotechnology , 130*, 316 - 319.

Wisselink, H. W., Toirkens, M. J., Franco Berriel, M., del Rosario, W., van Dijken, J. P., Pronk, J. T., et al. (2007). Xylose-metabolizing Saccharomyces cerevisiae strains overexpressing the TKL1 and *TAL1* genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase. *Applied Environmental Microbiology , 73*, 4881 - 4891.

Zaldivar, J. N. (2001). Fuel ethanol production from lignocellulose: a challenge for metabolic engineering and process integration. *Applied Microbial Biotechnology , 56*, 17 - 34.